

# SYSTEM AND METHOD FOR DATA PROCESSING

Publication number: JP9146878

Publication date: 1997-06-06

Inventor: GAI RIN GASURII; DANII MAABIN NIIRU; EDOWAADO JIYON SHIRUHA; SUTEIIBUN MAAKU SAABAA

Applicant: IBM

Classification:

- International: G06F13/40; G06F13/40; (IPC1-7): G06F13/36

- european: G06F13/40D5

Application number: JP19960263160 19961003

Priority number(s): US19950552034 19951102

Also published as:

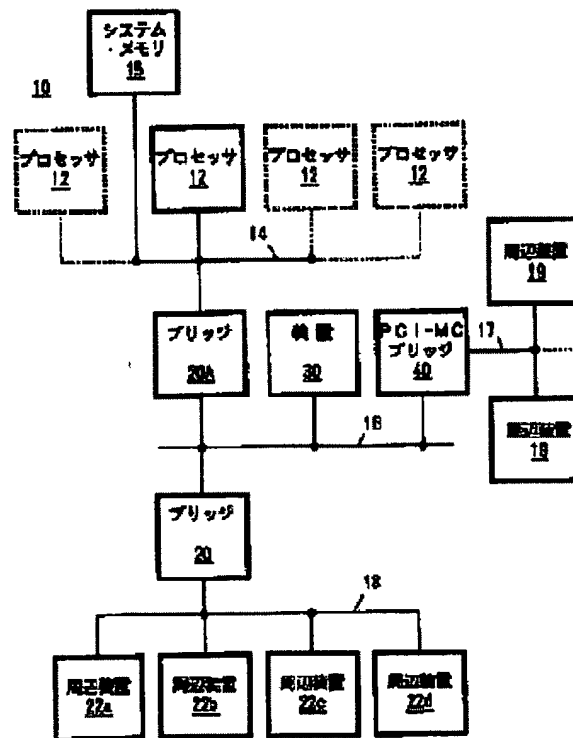
US5673399 (A)

Report a data error he

## Abstract of JP9146878

**PROBLEM TO BE SOLVED:** To efficiently handle many transactions coexisting in a host bridge in the both directions of inbound and outbound.

**SOLUTION:** A data processing system 10 includes a host processor 12, many peripheral devices 22 and one or more bridges 20. Each bridge 20 connects a first-order bus 14 (a system bus, for instance) and a second-order bus 16. A host bridge 20A includes an outbound data route, an inbound data route and a control mechanism. The outbound data route includes a queue buffer receiving transactions from the first-order bus 14 and storing the transactions in order and many parallel buffers storing reading response data and address information. Reading requests and writing requests are successively stored and reading response are stored in the many parallel buffers. The both of the inbound route and the outbound route in the host bridge 20A are considered in the activities in the both directions and are controlled by permitting or inhibiting a bypass transaction.



Data supplied from the esp@cenet database - Worldwide

(19) 日本国特許庁 (JP) (12) 特 許 公 報 (B2) (11) 特許番号  
特許第3579198号  
(P3579198)  
(45) 発行日 平成16年10月20日 (2004. 10. 20) (24) 登録日 平成16年7月23日 (2004. 7. 23)

(51) Int. Cl.<sup>7</sup> F I  
G 0 6 F 1 3 / 3 6 3 2 O A

請求項の数 14 (全 17 頁)

(21) 出願番号	特願平8-263160	(73) 特許権者	390009531
(22) 出願日	平成8年10月3日 (1996. 10. 3)		インターナショナル・ビジネス・マシー
(65) 公開番号	特開平9-146878		ズ・コーポレーション
(43) 公開日	平成9年6月6日 (1997. 6. 6)		I N T E R N A T I O N A L B U S I N
審査請求日	平成10年12月11日 (1998. 12. 11)		E S S M A S C H I N E S C O R P O
審判番号	不服2002-5030 (P2002-5030/J1)		R A T I O N
審判請求日	平成14年3月25日 (2002. 3. 25)		アメリカ合衆国10504 ニューヨーク
(31) 優先権主張番号	08/552034		州 アーモンク ニュー オーチャード
(32) 優先日	平成7年11月2日 (1995. 11. 2)		ロード
(33) 優先権主張国	米国 (US)	(74) 代理人	100086243
			弁理士 坂口 博
		(74) 代理人	100091568
			弁理士 市位 嘉宏

最終頁に続く

(54) 【発明の名称】 データ処理システム及びデータ処理方法

(57) 【特許請求の範囲】

【請求項 1】

少なくとも一つのプロセッサをメモリと接続するシステム・バスと、  
前記システム・バスとは異なるプロトコルを有する I/Oバスと、  
前記 I/Oバスに接続された、前記メモリおよび前記少なくとも一つのプロセッサとでデータ  
を共用する少なくとも一つの周辺装置と、  
前記 I/Oバスを前記システム・バスに相互接続し、アウトバウンド経路およびインバウ  
ンド経路を有し、前記少なくとも一つの周辺装置と前記メモリとの間の第一のデータ転送  
を、前記少なくとも一つのプロセッサと前記少なくとも一つの周辺装置との間の第二のデ  
ータ転送が完了する前に完了させることにより、トランザクションの順序を制御してデッ  
ドロック状態を防ぐブリッジとを含み、  
前記第一のデータ転送が完了するまで、前記少なくとも一つの周辺装置が処理活動を一時  
停止し、  
前記ブリッジが、  
前記少なくとも一つの周辺装置から前記メモリへの、前記第一のデータ転送に相当する要  
求を記憶する第一のバッファと、  
前記少なくとも一つのプロセッサから前記少なくとも一つの周辺装置への、前記第二のデ  
ータ転送に相当する要求を記憶する第二のバッファと、  
前記第二のバッファに対して並列である、前記第一のデータ転送に相当するデータを含む  
前記メモリからの返答を記憶する第三のバッファと、

10

20

前記メモリからの前記返答に対して前記少なくとも一つのプロセッサからの前記要求をバイパスさせ、前記少なくとも一つの周辺装置が処理活動を継続することができるようにする手段とを含むデータ処理システム。

【請求項 2】

前記第一のデータ転送が、前記少なくとも一つの周辺装置による前記メモリの直接メモリ・アクセスである請求項 1 記載のシステム。

【請求項 3】

前記第二のデータ転送が、前記少なくとも一つのプロセッサによるロードまたは記憶動作である請求項 2 記載のシステム。

【請求項 4】

前記ブリッジが、トランザクションを前記 I/Oバスから前記システム・バスに伝送するアウトバウンド経路を含む請求項 1 記載のシステム。

【請求項 5】

前記アウトバウンド経路および前記インバウンド経路が、  
前記ブリッジへの入力に接続された、前記アウトバウンド経路中でバッファへのトランザクションをゲートする第一のトランザクション・ルータ回路と、  
前記第一のトランザクション・ルータ回路に接続された、読み出し返答トランザクションを記憶する少なくとも一つの並列バッファと、  
前記第一のトランザクション・ルータ回路に接続された、読み出し要求トランザクションおよび書き込みトランザクションを記憶する少なくとも一つの直列バッファと、  
前記並列バッファおよび直列バッファの出力に接続された、状態機械の制御の下で前記アウトバウンド経路から出力される一つ以上のトランザクションを選択する第一のトランザクション選択回路とを含む請求項 4 記載のシステム。

【請求項 6】

前記ブリッジが、トランザクションを前記システム・バスから前記 I/Oバスに伝送するインバウンド経路をさらに含む請求項 5 記載のシステム。

【請求項 7】

前記インバウンド経路が、  
前記ブリッジへの入力に接続された、前記アウトバウンド経路中でバッファへのトランザクションをゲートする第二のトランザクション・ルータ回路と、  
前記第二のトランザクション・ルータ回路に接続された、読み出し返答トランザクションを記憶する少なくとも一つの並列バッファと、  
前記第二のトランザクション・ルータ回路に接続された、読み出し要求トランザクションおよび書き込みトランザクションを記憶する少なくとも一つの直列バッファと、  
前記並列バッファおよび直列バッファの出力に接続された、前記状態機械の制御の下で前記インバウンド経路から出力される一つ以上のトランザクションを選択する第二のトランザクション選択回路とを含む請求項 6 記載のシステム。

【請求項 8】

前記状態機械が、所定の論理構造にしたがって、前記アウトバウンド経路およびインバウンド経路中のトランザクションのゲート動作を制御する請求項 7 記載のシステム。

【請求項 9】

少なくとも一つのプロセッサをメモリと接続するシステム・バスと、  
前記システム・バスとは異なるプロトコルを有する I/Oバスと、  
前記 I/Oバスに接続された、それぞれが前記メモリおよび前記少なくとも一つのプロセッサとの間でデータを転送することができる複数の周辺装置と、  
前記 I/Oバスを前記システム・バスに相互接続し、アウトバウンド経路およびインバウンド経路を有し、前記 I/Oバス上の前記複数の周辺装置と、前記システム・バス上の前記少なくとも一つのプロセッサおよびメモリとの間の複数のデータ転送トランザクションを処理するブリッジとを含み、  
前記ブリッジは、

10

20

30

40

50

前記複数の周辺装置のうちの一つの周辺装置と前記メモリとの間の第一のデータ転送を、前記少なくとも一つのプロセッサと前記複数の周辺装置のうちの一つの周辺装置との間の第二のデータ転送が完了する前に完了させることにより、トランザクションの順序を制御してデッドロック状態を防ぎ、少なくとも一つの周辺装置から前記メモリへの、前記第一のデータ転送に相当する要求を記憶する第一のバッファと、前記少なくとも一つのプロセッサから前記少なくとも一つの周辺装置への、前記第二のデータ転送に相当する要求を記憶する第二のバッファと、前記第二のバッファに対して並列である、前記第一のデータ転送に相当するデータを含む前記メモリからの返答を記憶する第三のバッファと、前記メモリからの返答に対して前記少なくとも一つのプロセッサからの要求をバイパスさせ、少なくとも一つの周辺装置が処理活動を継続することができるようにすることを特徴とするデータ処理システム。 10

【請求項 10】

前記ブリッジが、前記複数のデータ転送トランザクションのうち、どの特定のデータ転送トランザクションを非順序的に処理するかを決定する状態機械を含む請求項 9 記載のシステム。

【請求項 11】

前記状態機械が、前記第一のデータ転送トランザクションによってアクセスされるデータを含む第一のアドレスが、前記後続のデータ転送トランザクションによってアクセスされるデータを含む第二のアドレスと一致するかどうかを判定するコンパレータを含む請求項 10 記載のシステム。 20

【請求項 12】

前記状態機械が、前記第一のアドレスと前記第二のアドレスとが一致したとき、前記後続のトランザクションが前記第一のデータ転送トランザクションをバイパスすることを防ぐため、前記後続のデータ転送トランザクションをゲートする手段をさらに含む請求項 11 記載のシステム。

【請求項 13】

少なくとも一つのプロセッサとメモリとをシステム・バスによって接続するステップと、前記システム・バスとは異なるプロトコルを有する I/O バスを設けるステップと、前記メモリおよび前記少なくとも一つのプロセッサとでデータを共用する複数の周辺装置を前記 I/O バスによって接続するステップと、  
前記 I/O バスを前記システム・バスにアウトバウンド経路およびインバウンド経路を有するブリッジにより相互接続し、前記複数の周辺装置の一つと前記メモリとの間の第一のデータ転送を、前記少なくとも一つのプロセッサと前記複数の周辺装置の前記一つとの間の第二のデータ転送が完了する前に完了させることにより、トランザクションの順序を制御してデッドロック状態を防ぐステップと、  
前記複数の周辺装置の前記一つを、前記第一のデータ転送が完了するまでデータ処理活動が一時停止される状態に維持するステップとを含み、  
前記ブリッジによる相互接続ステップが、  
前記複数の周辺装置の前記一つから前記メモリへの、前記第一のデータ転送に相当する要求を第一のバッファに記憶するステップと、  
前記少なくとも一つのプロセッサから前記複数の周辺装置の前記一つへの、前記第二のデータ転送に相当する要求を第二のバッファに記憶するステップと、  
前記第一のデータ転送に相当するデータを含む前記メモリからの返答を、前記第二のバッファに対して並列である第三のバッファに記憶するステップと、  
前記メモリからの前記返答をして前記少なくとも一つのプロセッサからの前記要求をバイパスさせて、前記複数の周辺装置の前記一つが処理活動を再開することができるようにするステップとを含む  
データ処理システムにおけるデータ処理方法。 30

【請求項 14】

前記第一のデータ転送が、前記複数の周辺装置の前記一つによる前記メモリの直接メモリ 40 50

・アクセスであり、前記第二のデータ転送が、前記少なくとも一つのプロセッサによるロードまたは記憶動作である請求項 13 記載の方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、データ処理システムに関し、より詳細には、ホストと周辺装置との間のインバウンドおよびアウトバウンドの読み出しおよび転記された書き込みを効率的に扱うための機構を含むデータ処理システムに関する。

【0002】

【従来の技術】

周辺部品インタコネクト (PCI) バス仕様のようなデータ処理システムに使用するための多くの標準バス・アーキテクチャは、同じバス・プロトコルを有する少なくとも二つのバスを接続するバス間ブリッジを横切るトランザクションのための順序づけ規則を含む。そのようなトランザクションの例は、ホスト・プロセッサからアウトバウンドに周辺装置に至る読み出し要求、周辺装置からインバウンドにホスト・プロセッサに至るデータを含む読み出し返答 (読み出し要求に対する) または書き込みコマンドがバス間ブリッジ中のバッファに転記されて、ホストが他の処理を進めることを許可する転記された書き込みである。本明細書中、「転記された書き込み」とは、開始側のバスにおいて書き込みが完了し、その書き込みが、行き先バスにおける将来の完了に備えてホスト・ブリッジ中に転記されているものをいう。現在のバス間アーキテクチャは、バス間ブリッジを横切る読み出しトランザクションを完了する前に、アウトバウンド (ホスト・プロセッサから離れる) およびインバウンド (ホスト・プロセッサに向かう) の両方向の書き込みバッファをフラッシュ消去することを要するか、特定のバス間順序づけ規則、例えば相互参照した特許出願 AA9-95-012 に前記し、PCI バス・アーキテクチャの PCI ローカル・バス仕様バージョン 2.1 に述べられたものを定義するかのいずれかである。バッファをフラッシュ消去する必要性は、プロセッサの読み出しアクセスを待機させ、ブリッジ・バッファ中に転記された書き込みがあるならば、そのようなアクセスが再試行されることにより、使用中のサーバの場合に性能の問題を生じさせるおそれがある。本明細書中、「再試行」とは、マスタによってアクセスされている標的装置がそのアクセスを認識するが、その装置が使用中であり、トランザクションを停止するという合図することに関する。そして、マスタは後で再びアクセスを試みる。

【0003】

一つの PCI バスを別の PCI バスに接続するバス間ブリッジを、本明細書中、「PCI-PCI ブリッジ」と呼ぶ。PCI-PCI ブリッジを横切るトランザクションのための順序づけ規則が、PCI ローカルバス仕様改訂 2.1 に包含されている。ホスト・プロセッサ、システム・バスまたは他のシステム相互接続を PCI バスに接続するブリッジを、本明細書中、「PCI ホスト・ブリッジ (PHB)」または単に「ホスト・ブリッジ」と呼ぶ。

【0004】

ホスト・ブリッジには、PCI-PCI ブリッジに要求されるものの他に独自の要件があるが、また、ホスト・ブリッジの場合、ホスト・ブリッジ・インタフェースの一次 (システム) 側の大部分のトランザクションのソースが I/O 装置よりもホスト・プロセッサによって開始されるため、PCI-PCI ブリッジと同じ制約のすべてがあるわけではない。

【0005】

いくつかの従来技術のブリッジ・アーキテクチャが、性能を改善し、処理能力の遅延を解消することを試みている。

【0006】

従来技術として、ネットワークを相互接続するマルチポート・ブリッジをもち、データを、読み出しが書き込みかを問わず、ブリッジ内での一様な扱いによって転送することを保

証する緩衝部品を含む高性能ブリッジのアーキテクチャがある。

【0007】

この技術は一般にブリッジの性能における改善に関するが、システムの性能を下げる読み出しトランザクションを完了する前にバッファをフラッシュ消去しなければならない問題に取り組んではいない。

【0008】

また、他の従来技術においては、多数の独立バッファ、メモリ・インタフェースおよび中央制御ユニットが接続されている共通のバスにより、メモリ・インタフェースがソース・ユーザからメッセージを受け、そのメッセージを選択されたバッファに記憶し、それらのバッファを連鎖する。制御装置が、それがメモリ・インタフェースから受けるコマンドに  
10 応答して、インバウンド・メッセージ・キューおよびアウトバウンド・メッセージ・キューを生成する。

【0009】

この技術は、高性能バス・ブリッジ・アーキテクチャに対して類似点をいくつか有するが、転記された書き込みのトランザクションに続く読み出しトランザクションを完了する前にバッファをフラッシュ消去する必要性の問題に対する解決方法を扱ってもしないし示唆  
してもいない。

【0010】

また、他の従来技術には、プロセッサから周辺装置に情報を読み出し、ネットワーク環境  
20 においてブリッジングを行うための多数のインバウンドおよびアウトバウンドのバッファがある。

【0011】

この技術は、バッファ・メモリを含むバス間ブリッジを有しているが、書き込みトランザクションと読み出しトランザクションとの間でバッファをフラッシュ消去する必要性を解消するバス間ブリッジにおける性能を改善するための機構を有していない。

【0012】

上述の従来技術のいずれも、書き込み動作に続く読み出し動作のためにバッファをフラッシュ消去する必要性を解消する性能の改善を教示してもしないし示唆してもしない。

【0013】

さらに、従来技術は、混在する多数のインバウンドおよびアウトバウンドのトランザクションを処理する問題を適切に扱っていない。加えて、従来技術は、異なるプロトコルを有するバスどうし、例えばシステム・バスとI/Oバスとをインタフェースさせて、システム干渉性定義域要件、I/Oバス要件、デッドロック回避および性能の観点を考慮しながら、トランザクション順序づけ要件を合わせなければならないホスト・ブリッジ独自の要件に対処していない。

【0014】

特に、システム・バスは通常、特定のトランザクションを、データを要求/提供している装置に対してタグ付けする能力を有するため、システム・バスをI/Oバスと相互接続する場合、特定の問題が生じる。しかし、I/Oバス（特にPCIバス）は、装置に対してトランザクションをタグ付けする能力を有さず、アドレスに基づいてデータにアクセスする  
40 だけであり、これがデータ・コヒーレンシーに関して問題を引き起こしかねない。すなわち、最初の装置が特定のアドレスからデータを要求すると、ブリッジがこの最初の装置に対して再試行を実施し、それに対し、後でデータを取りに戻ってくるように伝えることができる。そして、ブリッジは、特定のアドレスからデータを取得し、そのデータをブリッジ中のバッファなどに記憶して、第一の装置がそのデータを再び要求してくるのを待つ。第一の装置がデータを再び要求してくる前に、第二の装置が他のデータをこの特定のアドレスに書き込み、それを読み出そうとするかもしれない。第一の装置によって要求されるデータは今やブリッジ中のバッファの中にあり、通常のバス間ブリッジは、特定の装置  
50 に関してデータをタグ付けする能力を有さず、ただ、特定のアドレスからデータが要求されたことを知るだけであるため、そのデータは第二の装置に提供される。したがって、異

なる装置によって同じアドレスからデータが要求されるとき、第一の装置のためのデータが第二の装置に提供されることになるため、問題が起きることが理解されよう。このように、第二の装置が第一の装置からデータを「奪った」ことになり、それが誤った結果を招くおそれがある。装置が、別の装置に向けられた違うデータを取得してしまうことを防ぎ、かつ、トランザクションの順序外処理を可能にするブリッジが必要であることが理解されよう。

#### 【0015】

記憶トランザクションがブリッジの出力バッファ中で保留状態にある同じときに特定のI/O周辺装置がシステム・メモリに対して直接メモリ・アクセス(DMA)を試みるとき、もう一つの問題が生じる。これは、記憶がDMA動作の完了を待ち、DMAトランザクションが記憶の終了を待つところのデッドロック状態を引き起こす。より具体的には、多くのタイプのI/Oバスおよびそれらの対応する周辺装置は「再試行」能力をもたない。したがって、ある周辺装置がバスにアクセスするとき、その装置は、そのトランザクションが完了するまでバスの制御を維持する。例えば、周辺装置は、システム・バスに接続されたシステム・メモリからデータを要求するかもしれない。この場合、読み出し要求がブリッジのトランザクション待ち行列中に配される。続いて、メモリからの要求されたデータを含む読み出し返答動作がブリッジに提供される。しかし、その間に、システム・バスに接続されたプロセッサが記憶動作を実施して、先にDMAデータをシステム・メモリから要求していた周辺装置にデータを配置してしまうかもしれない。この記憶動作がブリッジ待ち行列中のDMAデータ返答トランザクションよりも先に起こるならば、デッドロック状態が起こる。すなわち、周辺装置が、DMAデータを受けるまで、他の動作を進めることができず、記憶トランザクションがブリッジ待ち行列中のDMAデータ返答トランザクションよりも先に起こるため、要求を出した周辺装置に向けられる記憶トランザクションは完了することができない。したがって、このデッドロック状態を解消し、データ処理システム中で改善されたデータ転送動作を可能にするブリッジが必要であることが理解されよう。

#### 【0016】

##### 【発明が解決しようとする課題】

したがって、本発明の目的は、システムの高い性能水準を維持しながらも、読み出し要求、読み出し返答、I/O書き込みおよび転記されたメモリ書き込みを効率的な方法で制御し、適切なホスト・ブリッジ・トランザクション順序づけを提供し、デッドロックを回避するための制御機構を有して、ホスト・ブリッジ中で多数の混在するトランザクションをインバウンドおよびアウトバウンドの両方向に効率的に扱うことにある。

#### 【0017】

##### 【課題を解決するための手段】

したがって、データ処理システムは、ホスト・プロセッサと、多数の周辺装置と、一つ以上のホスト・プロセッサ、周辺装置および相互接続された周辺装置をもつ他のホスト・プロセッサをネットワークにおけるように接続することができる一つ以上のブリッジとを含む。ホスト・ブリッジが、一次バス(例えばシステム・バス)と二次バス(例えばPCIバスのようなI/Oバス)とを接続する。ここで、理解しやすいよう、一次バスを、アウトバウンド・トランザクションのソースおよびインバウンド・トランザクションの行き先とみなし、二次バスを、アウトバウンド・トランザクションの行き先およびインバウンド・トランザクションのソースとみなす。具現態様に依存して、一次バス上の種々のトランザクションが二次バス上の他のトランザクションと同時に起こることもあるし起こらないこともある。各ホスト・ブリッジは、アウトバウンド・データ経路と、インバウンド・データ経路と、制御機構とを含む。アウトバウンド・データ経路は、トランザクションを一次バスからの受け取り順(または他の順序)に記憶するための待ち行列バッファであって、待ち行列バッファ中の要求が、読み出し要求と書き込み要求との間で混在していてもよいところの待ち行列バッファを含み、また、読み出し返答データおよびアドレス情報を記憶するための多数の他のバッファを含む。

## 【0018】

インバウンド経路はアウトバウンド経路と鏡像的に対照するものであり、読み出し要求および書き込み要求がトランザクション・バッファに記憶され、読み出し返答が多数の他のバッファに記憶される。ホスト・ブリッジ中のインバウンド経路およびアウトバウンド経路はいずれも、両方向への活動を考慮し、バイパス・トランザクションを許可または禁止する状態機械によって制御される。本明細書中、「バイパス」とは、あるトランザクションがその後続のトランザクションよりも先に待ち行列に加わったとしても、先行のトランザクションの前に後続のトランザクションを処理することが許可されることをいう。

## 【0019】

本発明の特徴は、インバウンドおよびアウトバウンドのトランザクション要求を、所定の状態機械の制御の下で、特定のタイプのトランザクションおよび定義された順序づけ規則に基づき、順次または非順次のいずれにも扱うことができることにある。

## 【0020】

本発明のもう一つの特徴は、バッファをフラッシュ消去することなく読み出しトランザクションが書き込みトランザクションの後に続くことを許可する機構により、ホスト・ブリッジ中のインバウンドおよびアウトバウンドのトランザクションを効率的に扱うことができることにある。

## 【0021】

本発明のもう一つの特徴は、種々のトランザクションがホスト・ブリッジを横切るとき、それらのトランザクションに関して適切なデータ整合性が維持されることにある。

## 【0022】

他のトランザクション、例えば読み出し要求は、開始側のバス上で完了する前に行き先バス上で完了して（データが取り出されて）いなければならない。例えば、ホスト・プロセッサは、I/Oバス上の周辺装置からデータを要求するかもしれない。そのデータは、トランザクションが完了したとみなされる前に、行き先バス上の周辺装置から取り出され、ホスト・ブリッジを介して開始側のバス上のホスト・プロセッサに提供されなければならない。本発明を特徴を用いると、転記された書き込みをフラッシュ消去することなく、トランザクション、例えば読み出しアクセスの処理を進めることができる。

## 【0023】

本明細書中、「転記された読み出し」とは、標的装置がアクセスを認識し、それが使用中であり、再試行制御信号によってアクセスを終了することを合図するものをいう。マスク（要求側の装置）は後で再びアクセスを試みる。そして、本発明のホスト・ブリッジが後続の処理に備えて読み出し要求を転記して、マスクが後で再びアクセスを試みる時、読み出されたデータを利用可能にできるようにする。ホスト・ブリッジが要求されたデータを取得する前にマスクが再び試行するならば、ブリッジは再び再試行を合図するが、アクセス要求の第二のコピーを緩衝記憶しない。ホスト・ブリッジの場合、プロセッサからの読み出し要求がタグ付けされてもよく、プロセッサが再試行されることを要しない。

## 【0024】

本発明のもう一つの特徴は、一般に、性能の改善および具現化しやすさのため、ロード動作（ホスト・プロセッサに対する読み出し動作）経路および記憶（ホスト・プロセッサに対する書き込み動作）が、周辺装置とシステム・メモリとの間の直接メモリ・アクセス（DMA）経路から独立していることにある。ロードおよび記憶動作は、ロードがホスト・プロセッサへのデータのロードを生じさせ、記憶がホスト・プロセッサから周辺装置へのデータの記憶を生じさせるというトランザクションの流れを考慮していることに注目すべきである。

## 【0025】

本発明のさらに別の特徴は、ホスト・ブリッジの周辺メモリ空間または周辺I/O空間（I/O出力バッファ）のいずれかへのロード要求が、同じホスト・ブリッジの周辺メモリ空間（メモリ出力バッファ）または周辺I/O空間のいずれかへの先行のロードの前には、両方のロードが同じアドレスに向かう場合、絶対にI/Oバスに通されてはならないと



いうことにある。アドレスが異なるとき、これは性能の改善を可能にする（通過を許す）が、一方または両方のプロセッサがその同じアドレスに書き込むとき、一方のプロセッサがもう一方のプロセッサのデータを奪う（すなわち無効データを得る）問題を回避する。

#### 【0026】

本発明のもう一つの特徴は、デッドロックを回避するため、DMA読み出し返答データが、ホスト・ブリッジ中に緩衝記憶された先行のロードおよび記憶トランザクションをバイパスすることを許可されることにある。これは、バス間ブリッジの場合とは異なる要件であり、性能を改善するためにホスト・ブリッジとPCIバスとの間の適切な相互作用のために求められる。

#### 【0027】

本発明のさらに別の特徴は、ホスト・ブリッジの周辺メモリ空間または周辺I/O空間のいずれかへのロードまたは記憶が、同じホスト・ブリッジの周辺メモリ空間または周辺I/O空間のいずれかへの先行の記憶の前には、絶対にI/Oバスに通されてはならないということにある。すなわち、一つのホスト・ブリッジによって生成された、I/Oバスに対する多数の記憶は、順序どおり維持されなければならない、ロードが記憶をバイパスしてはならない。

#### 【0028】

これらの技術は、同様に書き込みおよび読み出しトランザクションを転記するホスト・ブリッジ、一般にはバス間ブリッジに適用することができる。

#### 【0029】

前記の説明は、以下に記す本発明の詳細な説明をよりよく理解することができるよう、本発明の特徴および技術的利点を概説したものである。以下、本発明の請求の範囲の主題を構成する本発明のさらなる特徴および利点を記載する。

#### 【0030】

##### 【発明の実施の形態】

まず、図1を参照しながら、本発明を具現化するデータ処理システムを説明する。

#### 【0031】

システム10は、第一のバス14、例えばマルチプロセッサシステムによって使用することができるシステム・バス（例えばIBM社から販売されている6XXバス）に接続された少なくとも一つのホスト・プロセッサ12を含む。また、図1に点線で示す他のプロセッサ12がバス14に接続されていてもよい。

#### 【0032】

同様に、システム・バス14には、第一のバス14と第二のバス16とを接続するホスト・ブリッジ20aが接続されている。また、システム・メモリ15がシステム・バス14に接続されている。他の装置30、40が同様にバス16に接続されていてもよい。非限定的な一例として、装置40は、PCIバス（例えばバス16）を別のI/Oバス17、例えばIBM社から市販されているMicrochannelバス（MicrochannelはIBM社の商標である）に接続するための別のブリッジ・チップであってもよい。また、Microchannel（MC）周辺装置19が図示のとおりMCバス17に接続されている。したがって、MCデバイスは、PCI-MCブリッジ40およびホスト・ブリッジ20aを介してホスト・システム上の部品、例えばメモリ15、プロセッサ12などと通信することができる。本明細書にはMCを例として使用するが、本発明の範囲は、他のタイプのブリッジ・チップ40および周辺装置19をI/Oバス、例えばISA、ESA、VMEなどとともに使用することをとも考慮していることを記すべきである。さらには、別のブリッジ20がバス16および第三のバス18に接続されている。同様に、バス18には、多数のステーションまたは周辺装置22a、22b、22cおよび22dが接続されていてもよい。好ましい実施態様においては、ブリッジ20は、同一のプロトコルを有するバスどうし、すなわち、いずれもPCIバスであるバス16と18とを接続するバス間ブリッジである。ブリッジ20および20a以外、図1に関して上記に参照した各要素はすべて当該技術において周知であり、本明細書でさらに詳細に説明する必要は

10

20

30

40

50

ない。

【0033】

本発明の範囲を決して限定することなく、一例として、バス16および18はそれぞれ工業規格となったPCIバスであってもよい。

【0034】

次に、図2を参照しながら、ホスト・ブリッジ20aをさらに詳細に説明する。本発明の好ましい実施態様の説明に使用する際、「アウトバウンド」とは、プロセッサ12から離れるトランザクションに関していい、「インバウンド」とは、プロセッサ12に向かうトランザクションに関していう。

【0035】

ブリッジ20aは、二つのデータ経路、すなわちアウトバウンド経路202およびインバウンド経路204を有している。すべてのトランザクションは、インバウンドおよびアウトバウンドとも、状態機械206によって制御される。

【0036】

アウトバウンド・トランザクションがトランザクション・ルータ208によって受け取られると、トランザクション・フォーマットが審査されて、トランザクションが、読み出し要求（データの要求）、読み出し返答（データを提供する読み出し要求に対する応答）または書き込み要求（データを特定の装置に配置する要求）のいずれであるかが判定される。

【0037】

読み出し要求は、アドレスおよび制御情報のフォーマットを有している。読み出し返答は、アドレスと、制御情報と、すでに要求され、今や要求元に送られようとしているデータとを含むフォーマットを有している。

【0038】

書き込みトランザクション、例えば転記されたメモリ書き込み（PMW）またはI/O書き込み（I/O WR）は、アドレスと、制御情報と、選択されたアドレスに書き込まれるデータとを含む。

【0039】

例えば、現在トランザクション・ルータ208中で活動中であるトランザクションが読み出し要求であると仮定すると、そのトランザクションは、多数の直列バッファ210、212、214の一つに転送される。直列バッファ・エントリ210、212、214は、ブリッジ20a中に保留中であると予想されるトランザクションのスタックを扱うのに十分な長さであれば、いかなる適当な数のバッファ・エントリであってもよい。また、トランザクションが書き込みトランザクション、例えばI/O書き込みまたは転記されたメモリ書き込みであるならば、そのトランザクションは、直列バッファ210、212、214の一つに転送される。

【0040】

トランザクション・ルータ208中に保留中であるトランザクションが読み出し返答（RRP）であるならば、その読み出し返答トランザクションは、並列バッファ・エントリ216、218の一つに伝送される。

【0041】

いかなるときにも、トランザクション選択回路220によって選択され、二次バスに通されるトランザクションは、状態機械206によって制御される。これが、例えば、直列バッファ210、212もしくは214中に記憶されたトランザクションまたは並列バッファ216、218中に記憶された読み出し返答トランザクションに対する順序外アクセスを可能にする。

【0042】

インバウンド経路204は、各要素が逆の順序に接続されていることを除き、アウトバウンド経路202と同じ要素を含む。すなわち、トランザクション・ルータ222が二次バス16に接続されている（図1を参照）。アウトバウンド・トランザクション・ルータ2

10

20

30

40

50

08を参照しながら上記に説明したように、インバウンド・トランザクション・ルータ222は、状態機械206によって制御される。各トランザクションは、読み出し要求(RRQ)および書き込みトランザクション(I/O書き込みおよびPMW)の場合には、直列インバウンド・バッファ224、226および228の一つに経路指定される。読み出し返答トランザクションは、並列バッファ232、234に伝送される。インバウンド選択回路230が、状態機械206の制御の下で、直列バッファ224、226もしくは228または並列バッファ232もしくは234のいずれかからのトランザクションの伝送を制御する。

#### 【0043】

状態機械206は、図3および4に示す状態機械論理表によって定義されるプロトコルの下で動作して、アウトバウンドまたはインバウンドのいずれかまたは両方に流れる多数のトランザクションを制御する。

#### 【0044】

次に、図3および4を参照しながら、状態機械206を制御する論理表を説明する。図3の論理表の頂部にかけては、アウトバウンド・トランザクションを表す、便宜上「第一のトランザクション」と呼ぶトランザクションが並んでいる。図3の論理表の左側に沿っては、便宜上「第二のトランザクション」と呼ぶアウトバウンド・トランザクションが並んでいる。同様に、図4の論理表の頂部にかけては、インバウンド・トランザクションを表す、便宜上「第一のトランザクション」と呼ぶトランザクションが並んでいる。図4の論理表の左側に沿っては、便宜上「第二のトランザクション」と呼ぶインバウンド・トランザクションが並んでいる。

#### 【0045】

図3および4の表の項目に関して、「0」項目は、後続の(第二の)トランザクションをして先行の(第一の)トランザクションをバイパスさせてはならないことを意味する。「1」項目は、後続のトランザクションをブリッジに入らせて、必要ならば、トランザクションを進行させるため、先行のトランザクションをバイパスさせなければならず、さもなければ、デッドロックが起りうることを意味する。「0/1」項目は、後続のトランザクションをして先行のトランザクションをバイパスさせてもよいが、そうする必要もないことを意味する(順序づけ要件はない。つまり、どちらでもよく、性能および具現化コストの考慮に基づき、具現態様特定のことができる)。すなわち、特定の状況または具現態様に依存して、「0/1」は、ある場合にはバイパスを許可してもよく、他の場合にはバイパスを許可してはならない。

#### 【0046】

デッドロックを回避するため、DMA読み出し返答データには、ホスト・ブリッジ中に緩衝記憶された先行のロードおよび記憶トランザクションをバイパスすることを許可しなければならない。これは、バス間(例えばPCI-PCI)ブリッジの場合とは異なる要件であり、ホスト・ブリッジとPCIバスとの間の適切な相互作用のために必要である。これはまた、性能を改善することができる。

#### 【0047】

例えば、読み出し要求(RRQ)であるアウトバウンドの第一のトランザクションおよび読み出し返答(RRP)であるアウトバウンドの第二のトランザクションがあるならば、図3の論理表の縦列と横列との交差点が1を示す。1は、論理表のどの位置にあるかと、後続のトランザクションをして先行のトランザクションをバイパスさせなければならないことを示す。したがって、DMAデータをもつ読み出し返答(RRP)をして読み出し要求をバイパスさせなければならない。また、RRPをして書き込み動作、例えばPMWまたはI/O書き込みをもバイパスさせなければならないことが見てとれる。DMAデータ(RRP)動作をしてプロセッサのロードおよび記憶動作をバイパスさせることにより、デッドロック状態を引き起こすおそれのある前述の問題が解消される。例として、MC装置19がDMA要求によってシステム・メモリ15からデータを要求していると仮定する。すると、RRQトランザクションが直列バッファ224、226、228の一つに配置

される（これがインバウンド動作であるため）。この場合、RRQはバッファ228中にあると仮定される。DMAデータがシステム・メモリ15から戻される前に、プロセッサ12がMC装置19に対して書き込み動作を開始する。したがって、例えば、I/O書き込みトランザクションが直列バッファ210、212、214の一つに配置される（これがアウトバウンド・トランザクションであるため）。この場合、I/O書き込みがバッファ214中に配置されると仮定する。続いて、システム・メモリ15が、MC装置19からのRRQ（要求されたDMAデータを含む）に回答して、RRPを並列バッファ216または218の一方、例えばバッファ216に配置する。先に記したように、ある種のI/O装置、例えばMC装置19のプロトコルは、各トランザクションが完了するまで、バス17の制御を維持することをそれらに要求する。したがって、MC装置19が並列バッファ216中のDMAデータに仕えているためにバッファ214中のI/O書き込み動作が完了できず、潜在的なデッドロック状態が存在する。本発明は、バッファ216中のRRPトランザクション（DMAデータを含む）をしてバッファ214中のI/O書き込みトランザクションをバイパスさせて、MC装置19によるDMA要求を完了させることにより、このデッドロック状態の発生を防ぐ。当然、本発明は、他のデッドロック状態、例えばPMWまたはRRQトランザクションが直列バッファ214中にあり、バッファ216中のRRPと競合する状態の発生をも防止する。

#### 【0048】

0は、図3および4の論理表のどの位置にあるかと、バイパス・トランザクションの禁止を示す。したがって、第一のトランザクションとしてのアウトバウンド書き込み動作であって、第二のトランザクションとしてのアウトバウンド読み出し要求が後続するものが、その第二のトランザクションによってバイパスされてはならない。本明細書に使用する「バイパス」とは、第一の要求の完了の前に第二の要求を行き先バス上で完了させることをいう。しかし、第一のトランザクションが例えばアウトバウンド書き込みトランザクションであるならば、アウトバウンド読み出し返答トランザクションのバイパスを許可しなければならない。

#### 【0049】

図3および4の論理表の、バイパス・トランザクションの禁止を示す0が存在するすべての位置が書き込みトランザクションに関与することに注目すべきである。例えば、第二のトランザクションとしてのアウトバウンド読み出し要求は、第一のトランザクションとしてのアウトバウンド書き込みをバイパスすることはできない。図4から、インバウンド・トランザクションに関しても同じ原則が当てはまることが見てとれる。したがって、図4に示すように、インバウンド読み出し要求（RRQ）、読み出し返答（RRP）または書き込みトランザクション（PMW）が、第一のトランザクションとしての先行のインバウンド書き込み（PMW）をバイパスしてはならない。状態機械206の動作を制御する図3および4の論理表が、トランザクション・ルータ208および222ならびにトランザクション選択回路220および230それぞれにおけるトランザクションのゲート動作を制御する。当業者であれば、いかにして図3および4の表に表す論理を制御論理回路中に具現化し、状態機械206内に含めることができるかを理解するであろう。

#### 【0050】

PCIホスト・ブリッジ（PHB）が、I/OバスとしてのPCIバスを、独自の特徴を有するマイクロプロセッサ、例えばIBM社から市販されているPowerPC604マイクロプロセッサ（PowerPC およびPowerPC604はIBM社の商標である）を使用するコンピュータ・システム中のシステム・バスと接続する好ましい実施態様の場合には、取り組まなければならない問題点がいくつかある。一般に、性能を改善し、具現化を容易にするため、プロセッサ12に対するロード（RRQ）および記憶（PMW、I/O WR）経路は、周辺装置に対するDMA経路（例えばRRP）から独立しているべきである。また、図3および4の表における「0/1」項目に関して、これらのトランザクションは、バイパスを行ってもよいが、バイパスを行う必要もない。バイパスを行うか行わないかは、公平さの規則、性能の考慮または具現しやすさに基づいて決めるこ

とができる（以下の段落に記載する場合は除く）。

【0051】

PCIホスト・ブリッジの場合、ホスト・ブリッジの周辺メモリ空間または周辺I/O空間のいずれかに対するロード（例えばRRQトランザクション）は、同じホスト・ブリッジの周辺メモリ空間または周辺I/O空間のいずれかに対する先行のロードの前には、それらのロード両方が完全に同じアドレスに行く場合、絶対にI/Oバスに通してはならない。図3から、第一のトランザクションとしてのRRQ（ロード）動作と、第二のトランザクションとしてのRRQ（ロード）動作との交差点には「0/1」が存在することが見てとれる。この場合、すなわち標的のアドレスが同じではない場合、バイパスが許可される。しかし、状態機械206に含まれるコンパレータが、両アドレスが同じであると判定するならば、第二のRRQトランザクションが第一のRRQトランザクションをバイパスすることはできない。したがって、「0/1」（この場合）は、さらなる要因に依存して、例えばアドレス比較の結果に依存して、第二のトランザクションによる第一のトランザクションのバイパスが起こりうることを示す。これが、両アドレスが異なる場合には、性能の改善（追い越しを許可すること）を可能にするが、一方または両方の装置が同じアドレスでデータにアクセスしている場合には、一方の装置がもう一方の装置のデータを奪う（無効データを得る）問題を回避する。これらのトランザクションが異なるアドレスを有するならば、これらはバイパスすることができる。これが、両アドレスが異なる場合には、性能の改善を可能にするが、一方または両方の装置が同じアドレスに書き込む場合には、両方の装置が異なるアドレスでデータにアクセスしており、競合が存在しないため、一方の装置がもう一方の装置の読み出しデータを奪う（無効データを得る）問題を回避する。

【0052】

また、当業者であれば、いかにして状態機械206がコンパレータを含んで、要求側の装置（例えばプロセッサ12）によって標的にされるアドレスが同一であるかどうかを判定することができるかを理解するであろう。他に考慮すべきことには、PHBの周辺メモリ空間または周辺I/O空間は、その同じPHBの周辺メモリ空間または周辺I/O空間のいずれかに対する先行の記憶の前には、絶対にI/Oバスに通してはならない場合がある。すなわち、一つのPHBによって生成されるI/Oバスに対する多数の記憶は順序どおりに維持されなければならない、ロードが記憶を追い越してはならない。例えば、図3から、第二のトランザクションとしてのロード（RRQ）と記憶（PMWまたはI/O WR）との交差点には「0」が存在するため、第二のトランザクションが記憶を追い越すことができないことが見てとれる。さらに、図3の第一トランザクションと第二のトランザクションとの交差点の「0」から見てとれるように、第二のトランザクションとしての記憶（例えばPMW）が第一のトランザクションとしての別の記憶（PMWまたはI/O WR）を追い越すことはできない。

【0053】

すべての後続のインバウンド・トランザクションをして先行のインバウンドの転記されたメモリ書き込み（PMW）をバイパスさせてはならないことに注意すること。図4は、すべての第二のトランザクション（I/O WRは当てはまらないため、これは除く）と、第一のトランザクションとしてのPMWとの交差点に「0」を示す。また、図4の第一のトランザクションと第二のトランザクションとの交差点の「1」によって示されるように、第二のトランザクションとしての後続のインバウンド転記されたメモリ書き込みをして第一のトランザクションとしての先行の読み出し要求（RRQ）および読み出し返答（RRP）をバイパスさせなければならない。

【0054】

また、アウトバウンド・トランザクション対インバウンド・トランザクションおよびインバウンド・トランザクション対アウトバウンド・トランザクションを取り扱う表がないことに注目すること。これらの場合、トランザクションは反対方向に移動し、異なるアドレス空間に向かっているため、これらのトランザクションは独立しており、互いに対して順

序づけ要件をもたない。アウトバウンド・トランザクション対インバウンド・トランザクションまたはインバウンド・トランザクション対アウトバウンド・トランザクションが順序づけを要する場合には、それは、本発明の範囲を超えたソフトウェア・プロトコルによって扱われる。

【0055】

図4のいくつかの表項目が「na」であることが注目されよう。これらのケースでは、PowerPCマイクロプロセッサを用いるシステムの場合を取り扱う本発明の好ましい実施態様ではインバウンドI/O WRトランザクションが定義されていないため、「当てはまらない」のである。

【0056】

上記に挙げた考察事項および本明細書に引用例として含めるIBM社の特許出願第08/473,603号に挙げられた考察事項の他では、マイクロプロセッサ、例えばPowerPCマイクロプロセッサを用いるシステムにおけるPCIホスト・ブリッジに関するトランザクション順序づけ要件は、PCI-PCIブリッジの場合と同じである（PCIローカル・バス仕様バージョン2.1を参照）。

【0057】

読み出しアクセスを試み、再試行されるマスタは、アクセスが完了するまで再び定期的にアクセスを試行しなければならない。

【0058】

まとめとして、本発明の構成に関して以下の事項を開示する。

(1) 少なくとも一つのプロセッサをメモリと接続するシステム・バスと、前記システム・バスとは異なるプロトコルを有するI/Oバスと、前記I/Oバスに接続された、前記メモリおよび前記少なくとも一つのプロセッサとでデータを共用する少なくとも一つの周辺装置と、前記I/Oバスを前記システム・バスに相互接続し、前記少なくとも一つの周辺装置と前記メモリとの間の第一のデータ転送を、前記少なくとも一つのプロセッサと前記少なくとも一つの周辺装置との間の第二のデータ転送が完了する前に完了させることにより、トランザクションの順序を制御してデッドロック状態を防ぐブリッジと、を含むことを特徴とするデータ処理システム。

(2) 前記第一のデータ転送が完了するまで、前記少なくとも一つの周辺装置が処理活動を一時停止する上記(1)記載のシステム。

(3) 前記ブリッジが、前記少なくとも一つの周辺装置から前記メモリへの、前記第一のデータ転送に相当する要求を記憶する第一のバッファと、前記少なくとも一つのプロセッサから前記少なくとも一つの周辺装置への、前記第二のデータ転送に相当する要求を記憶する第二のバッファと、前記第二のバッファに対して並列である、前記第一のデータ転送に相当するデータを含む前記メモリからの返答を記憶する第三のバッファと、前記メモリからの前記返答をして前記少なくとも一つのプロセッサからの前記要求をバイパスさせて、前記少なくとも一つの周辺装置が処理活動を継続することができるようにする手段とを含む上記(2)記載のシステム。

(4) 前記第一のデータ転送が、前記少なくとも一つの周辺装置による前記メモリの直接メモリ・アクセスである上記(3)記載のシステム。

(5) 前記第二のデータ転送が、前記少なくとも一つのプロセッサによるロードまたは記憶動作である上記(4)記載のシステム。

(6) 前記ブリッジが、トランザクションを前記I/Oバスから前記システム・バスに伝送する第一のデータ経路を含む上記(1)記載のシステム。

(7) 前記第一のデータ経路および前記第二のデータ経路が、前記ブリッジへの入力に接続された、前記第一のデータ経路中でバッファへのトランザクションをゲートする第一のトランザクション・ルータ回路と、

10

20

30

40

50

前記第一のトランザクション・ルータ回路に接続された、読み出し返答トランザクションを記憶する少なくとも一つの並列バッファと、  
前記第一のトランザクション・ルータ回路に接続された、読み出し要求トランザクションおよび書き込みトランザクションを記憶する少なくとも一つの直列バッファと、  
前記並列バッファおよび直列バッファの出力に接続された、状態機械の制御の下で前記第一のデータ経路から出力される一つ以上のトランザクションを選択する第一のトランザクション選択回路とを含む上記(6)記載のシステム。

(8) 前記ブリッジが、トランザクションを前記システム・バスから前記I/Oバスに伝送する第二のデータ経路をさらに含む上記(7)記載のシステム。

(9) 前記第二のデータ経路が、  
前記ブリッジへの入力に接続された、前記第一のデータ経路中でバッファへのトランザクションをゲートする第二のトランザクション・ルータ回路と、  
前記第二のトランザクション・ルータ回路に接続された、読み出し返答トランザクションを記憶する少なくとも一つの並列バッファと、  
前記第二のトランザクション・ルータ回路に接続された、読み出し要求トランザクションおよび書き込みトランザクションを記憶する少なくとも一つの直列バッファと、  
前記並列バッファおよび直列バッファの出力に接続された、前記状態機械の制御の下で前記第二のデータ経路から出力される一つ以上のトランザクションを選択する第二のトランザクション選択回路とを含む上記(8)記載のシステム。

(10) 前記状態機械が、所定の論理構造にしたがって、前記第一および第二のデータ経路中のトランザクションのゲート動作を制御する上記(9)記載のシステム。

(11) 少なくとも一つのプロセッサをメモリと接続するシステム・バスと、  
前記システム・バスとは異なるプロトコルを有するI/Oバスと、  
前記I/Oバスに接続された、それぞれが前記メモリおよび前記少なくとも一つのプロセッサとの間でデータを転送することができる複数の周辺装置と、  
前記I/Oバスを前記システム・バスに相互接続し、前記I/Oバス上の前記複数の周辺装置と、前記システム・バス上の前記少なくとも一つのプロセッサおよびメモリとの間の複数のデータ転送トランザクションを処理するブリッジとを含み、  
第一のデータ転送トランザクションが、後続のデータ転送トランザクションとは異なるアドレスでデータにアクセスするとき、前記複数のデータ転送トランザクションが非順序的に処理されることを特徴とするデータ処理システム。

(12) 前記ブリッジが、前記複数のデータ転送トランザクションのうち、どの特定のデータ転送トランザクションを非順序的に処理するかを決定する状態機械を含む上記(11)記載のシステム。

(13) 前記状態機械が、前記第一のデータ転送トランザクションによってアクセスされるデータを含む第一のアドレスが、前記後続のデータ転送トランザクションによってアクセスされるデータを含む第二のアドレスと一致するかどうかを判定するコンパレータを含む上記(12)記載のシステム。

(14) 前記状態機械が、前記第一のアドレスと前記第二のアドレスとが一致したとき、前記後続のトランザクションが前記第一のデータ転送トランザクションをバイパスすることを防ぐため、前記後続のデータ転送トランザクションをゲートする手段をさらに含む上記(13)記載のシステム。

(15) 少なくとも一つのプロセッサとメモリとをシステム・バスによって接続するステップと、  
前記システム・バスとは異なるプロトコルを有するI/Oバスを設けるステップと、  
前記メモリおよび前記少なくとも一つのプロセッサとでデータを共用する複数の周辺装置を前記I/Oバスによって接続するステップと、  
前記I/Oバスを前記システム・バスに相互接続し、前記複数の周辺装置の一つと前記メモリとの間の第一のデータ転送を、前記少なくとも一つのプロセッサと前記複数の周辺装置の前記一つとの間の第二のデータ転送が完了する前に完了させることにより、トランザ

10

20

30

40

50

クシヨンの順序を制御してデッドロック状態を防ぐステップと、  
を含むことを特徴とするデータ処理システムにおけるデータ処理方法。

(16) 前記複数の周辺装置の前記一つを、前記第一のデータ転送が完了するまでデータ処理活動が一時停止される状態に維持するステップをさらに含む上記(15)記載の方法

。(17) 前記相互接続ステップが、  
前記複数の周辺装置の前記一つから前記メモリへの、前記第一のデータ転送に相当する要求を第一のバッファに記憶するステップと、  
前記少なくとも一つのプロセッサから前記複数の周辺装置の前記一つへの、前記第二のデータ転送に相当する要求を第二のバッファに記憶するステップと、  
前記第一のデータ転送に相当するデータを含む前記メモリからの返答を、前記第二のバッファに対して並列である第三のバッファに記憶するステップと、  
前記メモリからの前記返答をして前記少なくとも一つのプロセッサからの前記要求をバイパスさせて、前記複数の周辺装置の前記一つが処理活動を再開することができるようにするステップとを含む上記(16)記載の方法。

(18) 前記第一のデータ転送が、前記複数の周辺装置の前記一つによる前記メモリの直接メモリ・アクセスであり、前記第二のデータ転送が、前記少なくとも一つのプロセッサによるロードまたは記憶動作である上記(17)記載の方法。

#### 【図面の簡単な説明】

【図1】本発明を具現化するデータ処理システムのブロック図である。

【図2】図1のシステムによる本発明の特徴を表すホスト・ブリッジ（例えばPCIホスト・ブリッジ）のブロック図である。

【図3】図2のホスト・ブリッジにおけるトランザクションの制御を、アウトバウンド経路のトランザクションに関して示す論理表である。

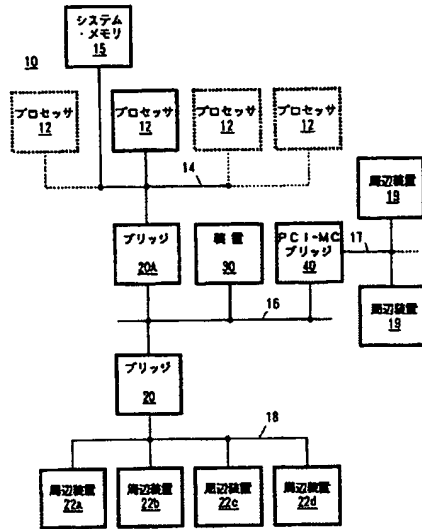
【図4】図2のホスト・ブリッジにおけるトランザクションの制御を、インバウンド経路のトランザクションに関して示す論理表である。

#### 【符号の説明】

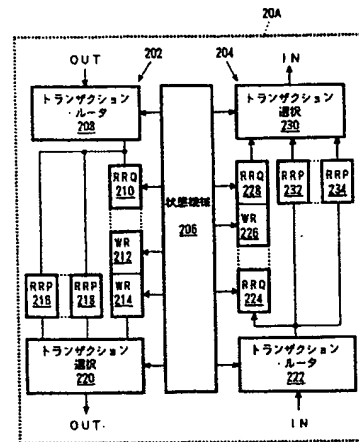
10	データ処理システム
12	プロセッサ
14	第一のバス
15	システム・メモリ
16	第二のバス
17	I/Oバス
18	第三のバス
19	周辺装置
20	ブリッジ
20A	ホスト・ブリッジ
22a、22b、22c、22d	周辺装置
30	装置
40	PCI-MCブリッジ



【図 1】



【図 2】



【図 3】

		アウトバウンド 第一のトランザクション			
		PMW (記憶)	RRQ (ロード)	IOWR (IO記憶)	RRP (DMAデータ)
第一の アウトバウンド トランザクション	PMW (記憶)	0	0/1	0	0/1
	RRQ (ロード)	0	0/1	0	0/1
	IOWR (IO記憶)	0	0/1	0	0/1
	RRP (DMAデータ)	1	1	1	0/1

定義：  
 アウトバウンド=ホスト・プロセッサから送れる  
 PMW=記憶されたメモリ番地  
 RRQ=読み出し要求  
 IOWR=I/O番地  
 RRP=読み出しデータ  
 DMAデータ=DMA読み込みデータ

【図 4】

		インバウンド 第一のトランザクション			
		PMW (DMAデータ)	RRQ (DMA読み出し)	IOWR (IO記憶)	RRP (ロードデータ)
第二の インバウンド トランザクション	PMW (DMAデータ)	0	1	na	1
	RRQ (DMA読み出し)	0	0/1	na	0/1
	IOWR (IO記憶)	na	na	na	na
	RRP (ロードデータ)	0	0/1	na	0/1

定義：  
 インバウンド=ホスト・プロセッサに向う  
 PMW=記憶されたメモリ番地  
 RRQ=読み出し要求  
 IOWR=I/O番地  
 RRP=読み出しデータ  
 DMAデータ=DMA読み込みデータ  
 DMA読み出し=DMA読み出し要求

## フロントページの続き

- (72)発明者 ガイ・リン・ガスリー  
アメリカ合衆国78726、 テキサス州オースチンキャラバー・ドライブ 11145
- (72)発明者 ダニー・マービン・ニール  
アメリカ合衆国78681、 テキサス州ラウンドロック ハイタワー・ドライブ 4604
- (72)発明者 エドワード・ジョン・シルハ  
アメリカ合衆国78759、 テキサス州オースチンビレネー・ドライブ 11509
- (72)発明者 スティーブン・マーク・サーバー  
アメリカ合衆国78717、 テキサス州オースチンエブライム・ロード 8308

## 合議体

審判長 下野 和行

審判官 東森 秀朋

審判官 矢島 伸一

- (56)参考文献 特開平5-314061 (JP, A)  
特開平7-210500 (JP, A)  
特開平3-262071 (JP, A)  
特開昭57-99843 (JP, A)